



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/661,245	09/12/2003	Bhashyam Ramesh	NCR 11092	8704
26890	7590	10/31/2007		
JAMES M. STOVER TERADATA CORPORATION 1700 SOUTH PATTERSON BLVD, WHQ5 DAYTON, OH 45479			EXAMINER CORRIELUS, JEAN M	
			ART UNIT 2162	PAPER NUMBER
			MAIL DATE 10/31/2007	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b> 10/661,245	<b>Applicant(s)</b> RAMESH ET AL.	
	<b>Examiner</b> Jean M. Corrielus	<b>Art Unit</b> 2162	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 17 August 2007.
- 2a) ☒ This action is **FINAL**.                      2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-12 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-12 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)                                | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                       | 5) <input type="checkbox"/> Notice of Informal Patent Application                       |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)<br>Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____  |

### **DETAILED ACTION**

1. This office action is in response to the amendment filed on August 18, 2007, in which claims 1-12 are presented for further examination.

#### ***Response to Arguments***

2. Applicant's arguments with respect to claims 1-12 have been considered but are moot in view of the new ground(s) of rejection necessitated by amendment.

#### ***Claim Rejections - 35 USC § 112***

3. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 1-12 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. It unclear how one having ordinary skill in the art would concluding that none of the unique n-gram are low frequency n-grams and have one or more pairs of high frequency simply by identifying a unique n-gram as recited in claims 4-5. It is also not clear how the applicant is defined low and high frequency pair. The specification does not provide any clarification of such. The variables R, TI...R, Ts, Tv, Tx, v, x, s, I, Tu, u, Tuv and Y set forth in claims 1, 6 and 10 are not defined in the specification to enable one having ordinary skill in the art to understand the use the invention as claimed. Applicant is advised to amend the claims by defining the variables set forth in the claims. Applicant is reminded that no new matter should be added.

***Remark***

4. In the office action mailed on October 13, 2005, the examiner has made reference to the variables R, Tl...R, Ts, Tv, Tx, v, x, s, I, Tu, u, Tuv and Y set forth in claims 1, 6 and 10 which are not defined in the specification to enable one having ordinary skill in the art to understand the use the invention as claimed. Variable S and V are still not defined in the claims. The rejection mailed on October 13, 2005 with respect the 112 rejection is still maintained.

5. The 101 rejection with respect to claims 1-3 and 6-12 set forth in the last office action is withdrawn.

6. The rejection 35 USC 103(a) with respect to claims 4-5 over Liddy et al., United States Patent No. 6,006,221 (hereinafter "Liddy") has withdrawn.

***Claim Rejections - 35 USC § 103***

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. Claims 4-5 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chandrasekar et al., (hereinafter "Chandrasekar") US Patent no. 6,578,032 in view of Perrow US Patent no. 7,284,009.

Art Unit: 2162

As to claim 4, Chandrasekar discloses a system for clustering character strings, wherein each character string comprises a word or a phrase. More important X discloses a system for clustering a first character string with another character string into one or more groups, when the first character string satisfies a predetermined degree of commonality with one or more character strings in each of these groups and when the first character string does not satisfy the predetermined level of commonality with another character string, another group is created, see col.2, lines 1-20. In particular, Chandrasekar discloses the claimed “identifying unique n-grams in each string” (col.7, lines 30-48); and ***“one or more pairs of high frequency n-grams from the strings are low frequency pairs and, in response, clustering each string with one or more clusters associated with low-frequency pairs of high frequency n-grams from that string”***(by calculating the frequency of the occurrence of the individual words and whole query, wherein the highest frequency words and queries are determined and the precise number of selected highest frequency "items" , the only one highest frequency item may be selected, where the subject item has a frequency score that is significantly higher than the second highest frequency item. If two or more highest frequency items are selected, it is determined whether the items have the same frequency score, if the scores are not the same, the highest frequency item may be selected as the topic, see col.). However, Chandrasekar does not explicitly disclose ***“none of the unique n-grams are low frequency n-grams and that”***. On the other hand, Perrow discloses the claimed “identifying unique n-grams in each string” (col.8, lines 25-30) ***“none of the unique n-grams are low frequency n-grams and that”*** (by simply deleting records from each n-gram table that have a very low frequency count (below a certain threshold; see col.8, lines 24-27); and ***“pairs of high frequency n-grams from the strings are low frequency pairs”*** (col.10 lines 60-64). Therefore, it

Art Unit: 2162

would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the teachings of the cited references, wherein the clustering technique as disclosed by Chandrasekar would incorporated the use of setting a guidelines wherein none of the unique n-grams are low frequency n-grams, in the same conventional manner as disclosed by Perrow. One having ordinary skill in the art would have found it motivated to use such a combination for the purpose of finding the common structures in the collection of strings.

As to claim 5, Chandrasekar discloses a system for clustering character strings, wherein each character string comprises a word or a phrase. More important X discloses a system for clustering a first character string with another character string into one or more groups, when the first character string satisfies a predetermined degree of commonality with one or more character strings in each of these groups and when the first character string does not satisfy the predetermined level of commonality with another character string, another group is created, see col.2, lines 1-20. In particular, Chandrasekar discloses the claimed "identifying unique n-grams in each string" (col.7, lines 30-48); and ***"one or more pairs of high frequency n-grams from the strings are low frequency pairs and, in response, associating that string with clusters associated with triples of n-grams including the pair"***(by calculating the frequency of the occurrence of the individual words and whole query, wherein the highest frequency words and queries are determined and the precise number of selected highest frequency "items" , the only one highest frequency item may be selected, where the subject item has a frequency score that is significantly higher than the second highest frequency item. If two or more highest frequency items are selected, it is determined whether the items have the same frequency score, if the

Art Unit: 2162

scores are not the same, the highest frequency item may be selected as the topic, see col.11, line 61-col.11, line 10); *“associating that string with clusters associated with triples of n-grams including the pair” (col.7, lines 48-50)*. However, Chandrasekar does not explicitly disclose *“none of the unique n-grams are low frequency n-grams and that”*. On the other hand, Perrow discloses the claimed “identifying unique n-grams in each string” (col.8, lines 25-30) *“none of the unique n-grams are low frequency n-grams and that”* (by simply deleting records from each n-gram table that have a very low frequency count (below a certain threshold; see col.8, lines 24-27); and *“pairs of high frequency n-grams from the strings are low frequency pairs” (col.10 lines 60-64)*. Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the teachings of the cited references, wherein the clustering technique as disclosed by Chandrasekar would incorporated the use of setting a guidelines wherein none of the unique n-grams are low frequency n-grams, in the same conventional manner as disclosed by Perrow. One having ordinary skill in the art would have found it motivated to use such a combination for the purpose of finding the common structures in the collection of strings.

***Conclusion***

9. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

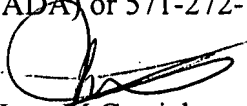
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jean M. Corrielus whose telephone number is (571) 272-4032. The examiner can normally be reached on 10 hours shift.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene can be reached on (571) 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.



Art Unit: 2162

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.



Jean M Corrielus  
Primary Examiner  
Art Unit 2162

October 25, 2007